MU-SPIN Ninth Annual Users' Conference

The Next Generation Web Servers

Dr. Mou-Liang Kung

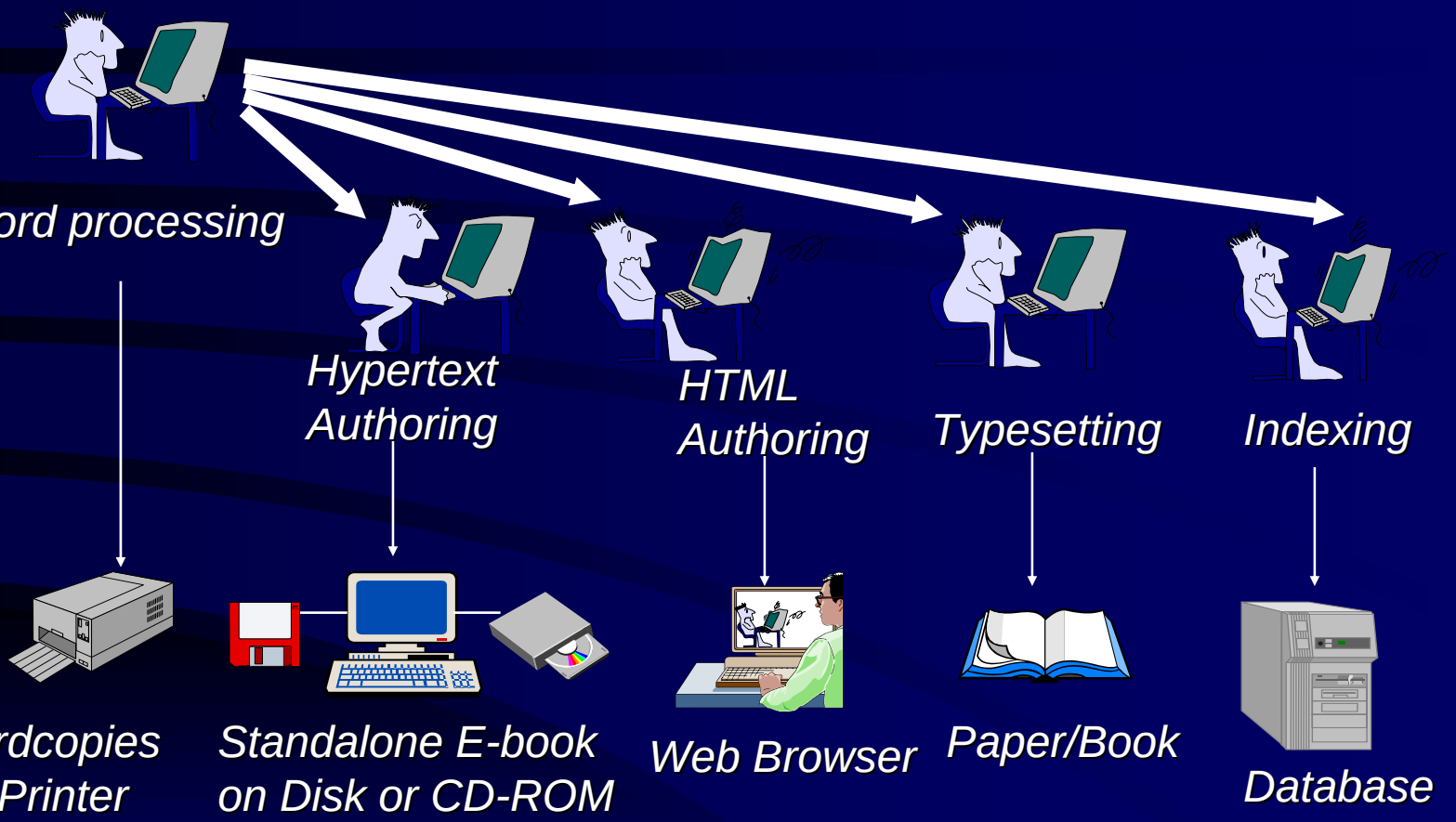
Department of Computer Science

Norfolk State University

Outline

- Functions of Web Servers
  - Document Publishing
  - Process Automation
- HTML Failed to deliver
- XML Comes to the Rescue
- XML-enabled Web Servers

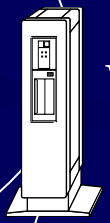
(Labor Intensive) Publishing Process



Web Publishing
- Electronic Flash Cards

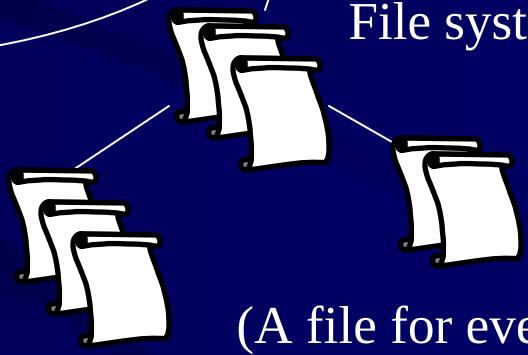
Web Browser

Internet



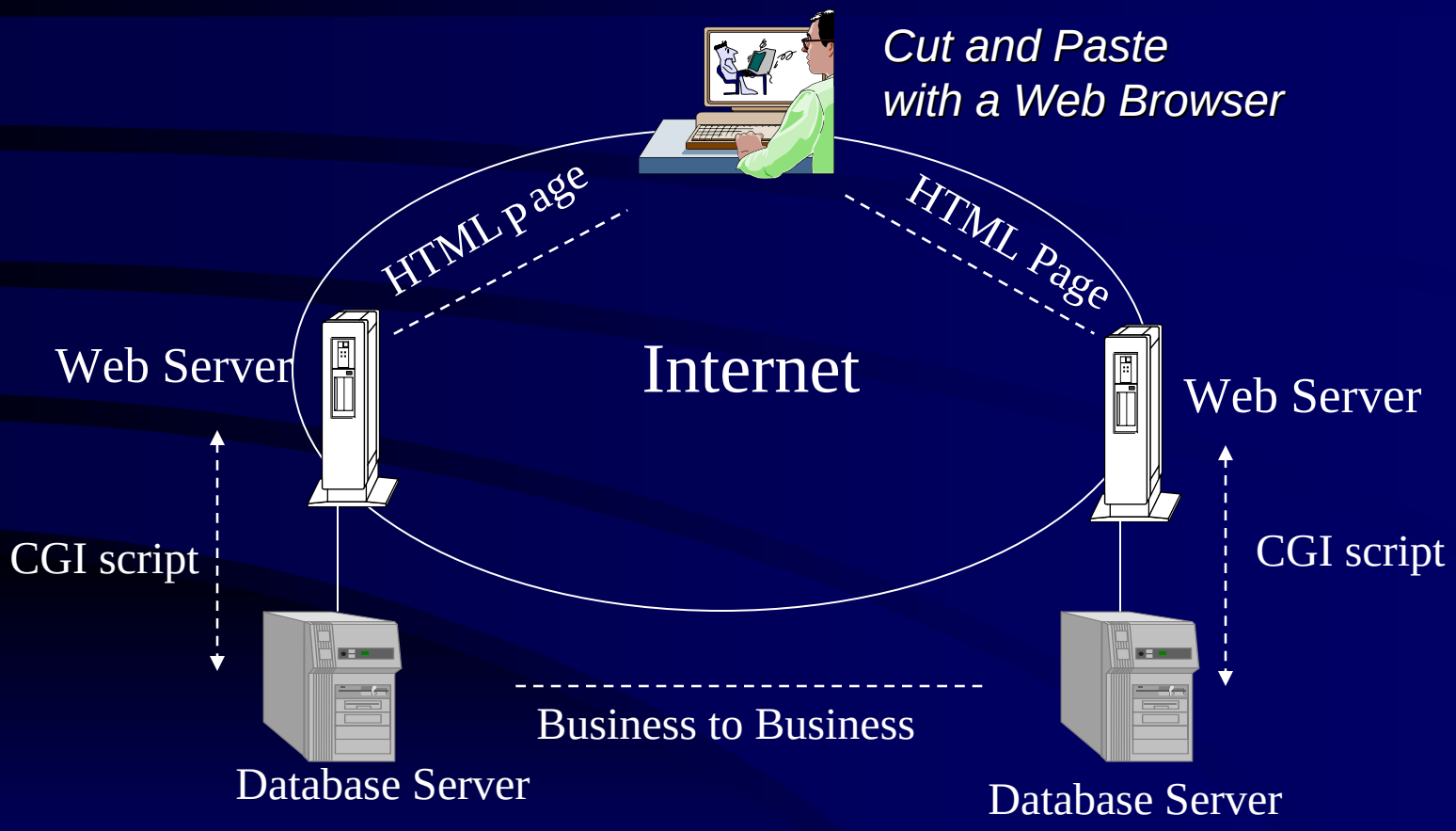
Web Server

File system



(A file for every Web page)

No Process Automation



Why HTML Failed to Deliver

- Fixed Content View (1 page, 1 view)
- Hard to implement Table of Content
- Primitive Hyperlinks: Unidirectional
  - Easily get lost in cyberspace
- Lack of structures
  - No intelligent search (full text search)
  - Pages can not be manipulated
  - Very poor for storage
  - No direct database query
  - No process automation

Why HTML Failed to Deliver

- Format tags (e.g. <b> for bold, for centering) mixed with structural tags (e.g. TITLE)
- Style sheet is hardcoded in the browsers*
- Fixed set of tags, No extensibility* (Documents created become locked to browsers, vice versa)

  (*Workaround through JAVA)
- Page-based displaying, difficult to convert large or sophisticated documents to HTML pages

HTML Is Incapable of Creating Applications that

- "require the Web client to mediate between two or more heterogeneous databases.
- attempt to distribute a significant proportion of the processing load from the Web server to the Web client.
- require the Web client to present different views of the same data to different users.
- in which intelligent Web agents attempt to tailor information discovery to the needs of individual users." (Bosak, [3])

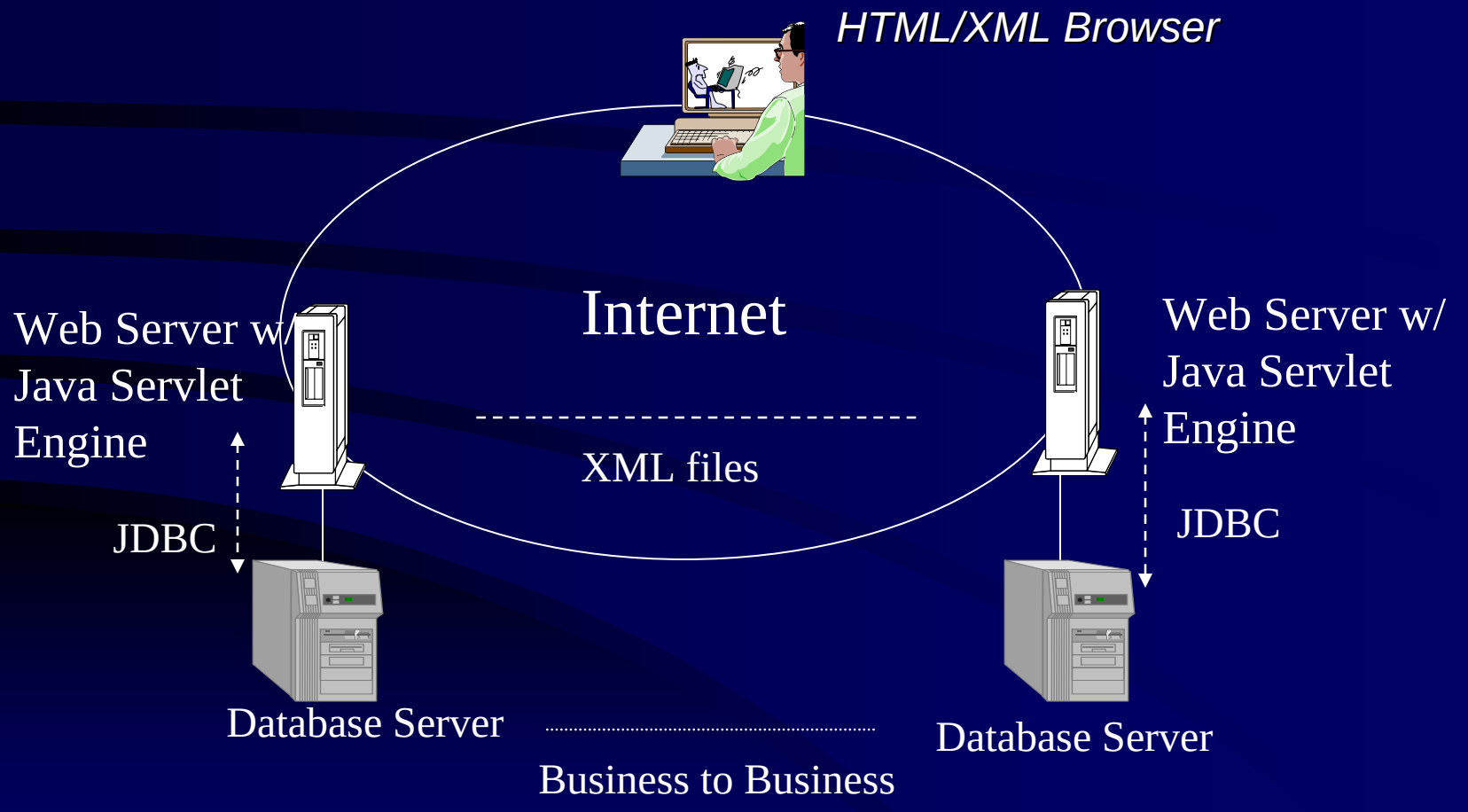
XML Comes to the Rescue

XML = eXtensible Markup Language,

a subset of SGML (ISO 8879:1986)

- Marks the structure and intent of a document, not the format
- Multiple Content Views for Access Control
- Reusable Document Components
- Can Be Converted to Multiple Formats on A Variety of Media for Any Application on Any System
- One-to-many, and Bi-directional hyperlinks

True Process Automation



XML Specifications

- XML
  - Well-formed
  - Valid, Conformed to Define Document Types (DTD)
- XLL, eXtensible Link Language:
  - Xlink, and
  - XPointer
- XSL, eXtensible Style Language

A Sample Memorandum

Memorandum

To: *Computer Science Faculty*
From: *Jim Kung*
Date: *September 18, 1999*
Subject: *Weekly Announcement*

We will have a Departmental picnic next Saturday. Don't forget to bring a cover dish.

Please attend next seminar on September 25, 1999 by Dr. S. Shah

As a Well-Formed XML

```
<?xml  version="1.0"  standalone="yes" ?>
<memo>_____Memorandum
        <head>
                <receiver>Computer Science Faculty </receiver>
                <sender>          Jim Kung </sender>
                <date>September 18, 1999</date>
                <subject>Weekly Announcement </subject>
        </head>
        <body>
                <para>
                We will have a Departmental picnic this Saturday.
        Don't forget to bring a covered dish.</para>
                <para>
                Please attend the next seminar on <date>September 25,
                1999 </date>by Dr. S. Shah.</para>
        </body>
</memo>
```

Memo DTD

```
<!--    This is a very simple DTD for Memoranda.-->
<!DOCTYPE    memo [
<!ELEMENT    memo    (head, body)                              >
<!ELEMENT    head    (sender, receiver, date, heading)         >
<!ELEMENT    (sender | receiver | heading)    (#PCDATA)        >
<!ELEMENT    date    (#PCDATA)                                 >
<!ELEMENT    body    (para+)                                   >
<!ELEMENT    para    (#PCDATA | date)                          >
]>
```

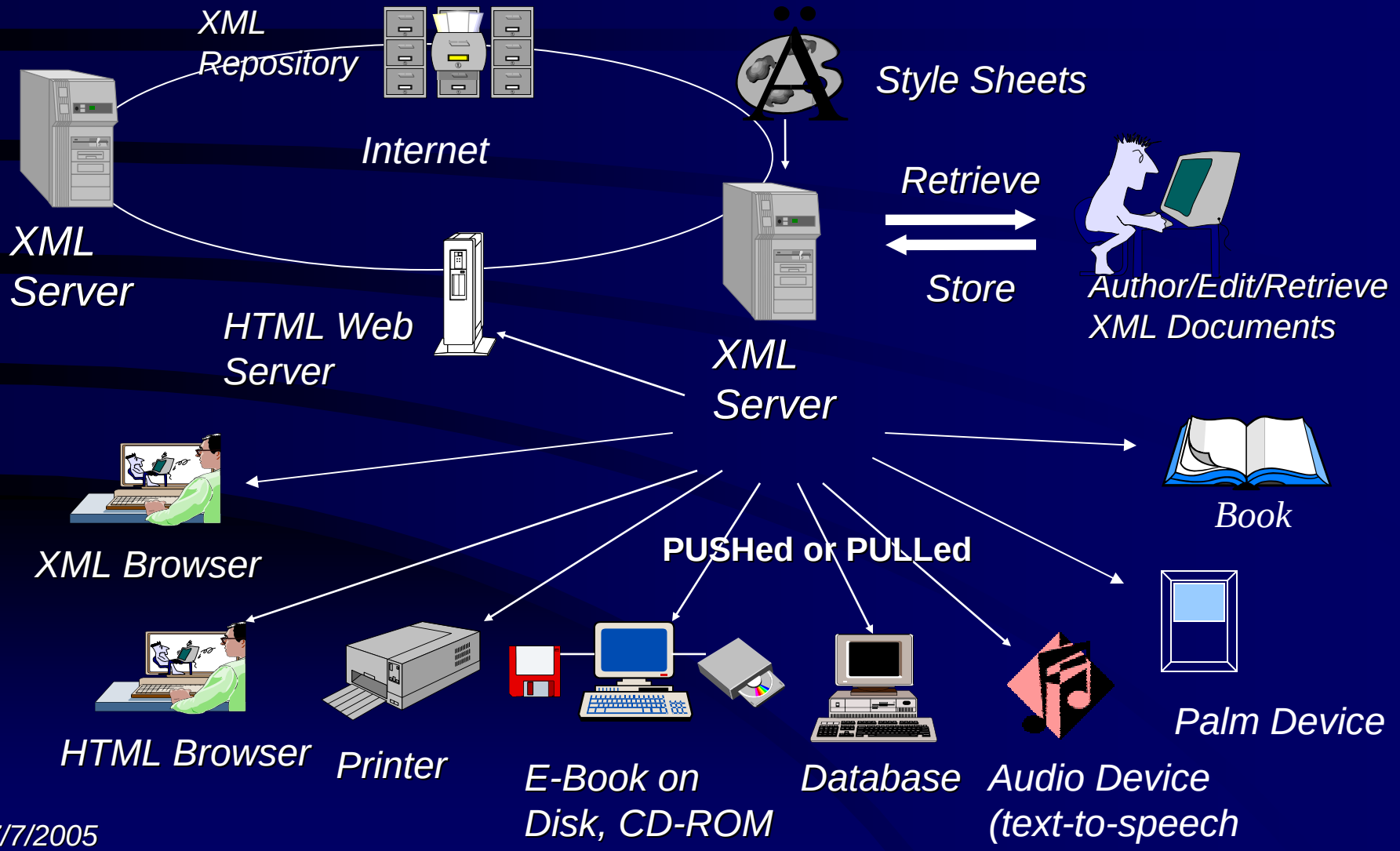
Appear in Sequence

Parse Character Data (Plain Text)

+ = Required & Repeatable
** = Optional and Repeatable*
? = Optional

The Next Generation Web Server

Using XML Technology



Enable XML on a Web Server

After Documents are Created in XML:

1. Distribute XML Documents to XML-Capable Browsers

For the HTML-only Browsers:

2. Downward Convert XML to HTML on the Fly Using Perl or Other Scripts
3. Downward Convert XML to HTML on the Fly Using Java Servlets

XML and Java

- Both XML and Java Support Unicode
- Java Supports Package Structure, Dynamic Class Loading, and JavaBeans API
- Both Designed for Distributed Systems

Anatomy of An XML Server

XML Browser

Internet

XML Server

XML Server

Communication Protocols

Document Handlers

XML Services:
Parser, XSL Engine, XQL

Data Access Objects

Other Java Services

Java VM

Database Server

XML Server Using Java Servlets

- Parser with DOM Support (e.g. IBM XML4J)

  where DOM (Document Object Model) is a set of APIs to access and modify the content, structure, and style of an XML or HTML documents.
- XSL (Extensible Stylesheet Language) engine (e.g. XSL:P)
- Java Servlet Engine (e.g. Apache JServ)
- XML->HTML Servlet (e.g. Apache Cocoon)

The Web Applications of the Future Are Here ...

- More Powerful Browsers (Jumbo, BSML Browser)
- Web Surfing on PDA, Digital Phones, Pagers and WAP (Wireless Application Protocol)
- XML will be incorporated in Oracle (8i), Sybase Database engines for easy database exchanges.
- For Internet Commerce, XML-EDI will merge ANSI X12 and U.N. EDIFACT into a single standard

XML Is Still Evolving
(You can get on the bandwagon)

- XSL and XLL are yet to be finalized
- XML Data Type support (v2.0?)
- XSchema: More Powerful Way to Specify Document Structure than DTD
- The Next Killer (XML) Application? Better Search Engine Using RDF Perhaps
- New XML Editors, XML Browsers: Most Current Ones are Prototypes
- Powerful XML Browsers Can Be Built on Top of GECKO

XML Tools

XML Editors

- **XML Pro**, http://www.vervet.com
- **CLIP**, http://xml.t2000.co.kr/product/intro.html
- **Xed**, http://www.xml.com/xml/pub/XED
- **Xmetal**, http://www.sq.com
- **Microsoft XML Notepad**, http://www.microsoft.com
- **Xeena**, http://www.alphaworks.ibm.com

Word Processors with XML Capability

- Corel Word Perfect 9
- Microsoft Office 2000

XML Tools

XML Browsers

- **MS IE 5.0,** http://www.microsoft.com
- **Gecko**, http://developer.netscape.com/software/communicator/ngl/index.html?cp=dev09fg01
- **Indelv,** http://www.indelv.com/browser.
- **Amaya,** http://www.w3.org/Amaya/
- **JUMBO**, a prototype GUI browser/editor/search/rendering tool, http://www.vsms.nottingham.ac.uk/vsms/java/jumbo/w3_blurb.html

Parsers

- **AElfred** from MicroStar, http://www.microstar.com
- **DXP** from DataChannel, http://www.datachannel.com/products/xdk/DXP
- **xml4j** from IBM (probably the most versatile), htttp://www.alphaworks.ibm.com/formula/xml
- **MSXML** from Microsoft http://www.microsoft.com/standards/xml/xmlparse.htm

XML Tools

XSL Style Sheet Editors

- **MSXSL** (free) from Microsoft takes an XML document and let user apply rules to produce an HTML document for non-XML Web browsers. http://www.microsoft.com
- **XML Styler** (free) from ArborText is probably the best one at present. http://www.arbortext.com/XML_Styler/xml_styler.html

DTD Editors

- **ezDTD (by Duncan Chen)** One can create or edit DTD
  http://www.geocities.com/SiliconValley/Haven/2638/ezdtd.zip
- **DTDGenerator** (written in Java by Michael Kay of ICL): Users input a well-formed XML, it will output a DTD
  http://home.iclweb.com/icl2/mhkay/dtdgen.html

XML Tools

XML Document Servers

- **DynaWeb**, INSO Corporation, www.inso.com
- **XMLServer**,BlueStone Software, www.bluestone.com
- **Content Management Suite**, Poet Software, www.poet.com
- **eXcelon**, Object Design Inc.,www.odi.com
- **Content management**, Ardent Software,www.ardentsoftware.com

Annotated References

[1] http://www-tei.uic.edu/orgs/tei/sgml/teip3sg/index.html

A Gentle Introduction to SGML, A Must-read for SGML Beginners

[2] http://www.sq.com/sgmlinfo/printintr.html

An SGML PRIMER

[3] http://sunsite.unc.edu/pub/sun-info/standards/xml/why/xmlapps.htm

XML, JAVA, and the Future of the Web, a Must-read

[4] http://www.csclub.uwaterloo.ca/u/relander/Academic/XML/xml_mw.html

XML: The New Markup Wave, a Must-read

[5] http://www.ucc.ie/xml/faq.html

XML Frequently Asked Questions, a Must-Read

[6] http://www.sil.org/sgml/sgmlnew.html

Latest News on SGML (a lot of XML news), a Must-read

[7] http://www.w3.org/XML/Activity

SGML, XML, and Structured Document Interchange

[8] http://www.w3.org/TR/1998/REC-xml-19980210.html

XML-- A W3C Recommendation

[9] http://www.w3.org/Press/1998/XML10-REC-fact

An XML Fact Sheet

[10] http://webreview.com/97/05/16/feature/xmldim.html

A four-part introduction on XML

[11] http://webreview.com/97/04/11/feature/

The Web is Ruined, and I Ruined it, "a kid named Marc Andreessen came up with the idea of the <IMG> tag, and the Web was both born and destroyed at that moment,…", where HTML ends and XML starts

[12] http://www.oasis-open.org/cover/xml.html#xmlSoftware

http://www.oasis-open.org/cover/publicSW.html

http://www.stud.ifi.uio.no/~larsga/linker/XMLtools.html#P_xt

Useful XML Software Pages

[13] http://www.oasis-open.org/cover/

Robin Cover's XML Home Page, A Must-Read